To: Commissioner Kenneth Kimmell and Assistant Commissioner Bethany Card, Massachusetts Department of Environmental Protection
From: N. LeRoy Poff, Ph.D.
Department of Biology, Colorado State University, Fort Collins CO 80523
RE: Comments on TRC Report
Date: 17 June 2012

I was asked by the Massachusetts Department of Environmental Protection (MDEP) to provide written comments on the several issues raised in response to the "TRC Report" or "rebuttal document" which was prepared for the Massachusetts Water Works Association (MWWA) and submitted to the Massachusetts Department of Environmental Protection (MDEP). These issues are:

1. A review of those portions of the USGS report http://pubs.usgs.gov/sir/2011/5193/ which are questioned by the rebuttal document. (USGS SIR 2011-5193): Citation: Armstrong, D.S., Richards, T.A., and Levin, S.B., 2011, Factors influencing riverine fish assemblages in Massachusetts: U.S. Geological Survey Scientific-Investigations Report 2011-5193, 58p;

2. A review of the rebuttal document (attached); and,

3. A summary of the validity of the rebuttal document as it relates to use of the USGS report to establish policies regarding water withdrawals and stream.

The TRC Rebuttal Document and My Process for Developing Comments

In preparing these comments, I **first** read the TRC Rebuttal Document (dated Feb. 2012) and the MWWA Comments document (dated Feb. 3, 2012), as provided me by MDEP. These documents make serious assertions about the scientific validity of the modeling approach used in the Armstrong et al. (2011) report. I listed what I viewed to be the most serious challenges levied in the TRC Rebuttal Document and the accompanying MWWA Comments against the scientific credibility of the analysis and findings in the USGS Report. **Second**, I carefully read the USGS Report (Armstrong et al. 2011) to evaluate the potential validity of those challenges. **Third**, I made a list of "questions for clarification" that I wished to put to the authors of the USGS Report; these questions focused specifically on methodological issues in model development that were not entirely clear to me from a reading of the USGS report. I had a brief conversation with them on May 27, 2012. The goal was to ask these individuals to clarify some of my questions on the underlying model framework and its application to the statewide scale. **Fourth**, I revisited the TRC Rebuttal Document and USGS documents and wrote my summary comments.

In this process, I focused on those TRC assertions that could be assessed by a careful reading of the USGS report, i.e., I did not address every detailed criticism, as many of these could not be "adjudicated" from a close reading of the USGS report (where not every detail was included) and would require a "cross-examination" of all parties. I adopted the perspective that if the most significant charges were found to be meritorious, then the smaller, more detailed criticisms would merit further review.

As a general comment, I would say I found the issues raised in the TRC Report and the MWWA document to be concerning, but the authoritative nature of the assertions and proposed alternative interpretations offered were weak. Most challenges are assertions that lack standard scientific authority such as scientific publications or regional experts. Most conclusions are based on what appear to be very selective examples that lack demonstrated statistical support.

By contrast, the Armstrong, Richards and Levin (2011) report is of high scientific quality, being well referenced and obviously peer-reviewed. This is a traditional scientific publication (which of course does not necessarily imply that it is beyond challenge.)

Thus, the TRC challenges deserve to be closely scrutinized for logic and credibility.

Background on the USGS Report and the FFRA Model

Armstrong et al. (2011) developed a statistical model relating the relative abundance of fluvial fish species (FFRA) to several environmental metrics across 669 sampling sites in the state of Massachusetts. A primary finding of the analysis was the establishment of a quantitative relationship between FFRA and intensity of groundwater pumping that depletes streamflow levels in August.

Two modeling approaches were adopted: a quantile regression (QR) model and a general linear model (GLM) with environmental covariates. Both models agree that FFRA declines monotonically with increasing August groundwater withdrawals (AUGgwWp). The GLM model shows that, at a statewide level across a gradient of flow alteration (i.e., across sites that vary in groundwater depletion intensity) that FFRA declines, on average, by 0.9% for each 1.0% increase in August groundwater depletion (Table 8, USGS Report). This finding is consistent with ecological theory and with a large published literature on ecological responses to hydrologic alteration (see citations in Armstrong et al. 2011).

The QR model clearly shows that FFRA to decline steeply with AUGgwWp. This modeling approach is appropriate for illustrating the general relationship between the response variable (FFRA) and the "limiting factor" of flow alteration. In some cases, as acknowledged in the USGS report, conditions other than flow alteration can be limiting (e.g., water quality, dams), which is why near-zero values of FFRA occur at localities that have little to no flow alteration.

The GLM model explains variation in *average* FFRA across 669 sampled stream localities from across Massachusetts. This model is presented in Equation 6 on p. 44 of the USGS report. The response variable in the model is catch per unit effort (CPUE), i.e., the number of fish captured at a locality *adjusted* by the effort (time) expended to capture those fish. CPUE is a <u>standardized</u> <u>measure</u> of fish abundance that allows inter-site comparisons to be made. The explanatory (or predictor) variables in the GLM model are: local channel slope (CHSLP), percent alteration of August median flow from groundwater withdrawals (AUGgsWp), percent wetland in a 240 meter buffer (pBWet) and percent impervious cover in the contributing drainage area (IC). At a statewide scale, this GLM explained about 18% of the among-site variation in standardized fluvial fish relative abundance, with all four explanatory variables being statististically significant at the 5% or lower level (Table 8, p. 41, USGS report).

The 669 sites selected for the FFRA models were screened according to a set of criteria (described on pp. 6-7, USGS report), which included ensuring that only sites that were efficiently sampled via backpack and/or barge electroshocking were included. Thus, the results and inferences from the model apply to small streams that are <u>wadeable</u> and can be electroshocked with backpack and barge shocking.

As mentioned on p. 49 of the USGS report, the GLM model could be used in the future to develop expectations of average FFRA for the 1,429 subbasins delineated in the State of Massachusetts. This process appears to have occurred after the publication of the USGS report, as presented in Figure 5 in the TRC report (Oct. 18, 2011 draft biological category map).

<u>The TRC Rebuttal Document</u>

There are 5 fundamental challenges to the USGS/MDFW model made in the TRC Rebuttal Document:

(1) The TRC Rebuttal Document asserts that the USGS/MDFW analysis is confounded by a "natural" West-East gradient in fluvial fish species presence. Thus, the claim is that USGS/MDFW model conflates water withdrawal's effects on fluvial fish abundance with a natural geographic gradient.

No convincing evidence of a West-East gradient in fluvial fish abundance is presented by TRC. The Figure 1 in the TRC Report shows some distribution of fish species but the source of this map and its accuracy are unclear and undocumented. While there may well be some West to East gradient in expected fish species occurrence, the relevant output variable for the USGS/MDFW model is FFRA, and the <u>abundance</u> of fluvial fish is not shown in the TRC Document. A more compelling demonstration would be to show that FFRA shows a geographic bias based on mapped output from the USGS/MDFW database.

From the model description (Equation 6, p. 44, USGS report), it is clear that several factors change along a West to East gradient in Massachusetts, such as impervious cover (Figure 5) and August withdrawals (Figure 6). The GLM model accounts for geographic variation in FFRA by using meaningful watershed and habitat predictor variables that are able to detect geographic gradients and that can be used to explain variation in fluvial fish relative abundance. These explanatory variables were "selected" based on statistical criteria of "best fit" or "most variance explained." This is a standard, <u>objective</u> method of variable selection. Further these variables (impervious cover, water withdrawal) have a "mechanistic" basis for explaining variation in fish performance and abundance. The coefficients shown in Table 8 indicate that any variation can generate spatially disjunct patterns of expected FFRA that would not occur with a simple "longitude" type variable, i.e., streams in eastern Massachusetts could have higher FFRA if they were less urbanized and had a lot of wetland buffer. (This is somewhat reflected in Figure 5 in the TRC Rebuttal Document, although that map does not include explanatory variables of wetland buffer and channel slope).

In sum, the geographic gradient of concern in the TRC Document appears to be captured reasonably within the model structure.

(2) TRC claims the FFRA model development was biased in that it included sites that were inefficiently sampled by electroshocking and thus inferences are questionable;

The sites included in the FFRA model were reported (USGS Report, p. 6) to be screened for efficiency of sampling using both backpack electroshocking (shallow water) and barge electoshocking sampling (deeper water in wadeable streams).

Streams with habitats too deep to sample were not included in the construction of the FFRA. Thus the assertion of biased sampling as a serious flaw in the FFRA does not appear to be supported by information available in the USGS report. The USGS/MFDW approach is, by definition, restricted to wadeable streams.

(3) Young-of-year fish were selectively excluded from FFRA models either or the basis of size or season of sampling, such that the model inputs of fluvial fish abundance are unreliable

The TRC Rebuttal Document asserts that small (< 40 mm) YOY fish were selectively eliminated from samples such that underestimation of fluvial fish abundance occurred, and that variation in sampling date is a source of potential bias in the estimation of RFFA. The USGS Report clearly states (p. 7) how YOY fish were included in the analysis. There is insufficient information presented in the USGS report to fully judge this TRC claim; however, the TRC Report fails to establish that there was any substantial geographic bias in the application of YOY screening or in date of sampling that would systematically (and geographically) bias the FFRA estimates statewide.

(4) The "predicted" numbers of fluvial fish at sites with existing data are inaccurate; therefore, the FFRA model is flawed and not useful for regulatory purposes;

This is perhaps the most important claim in the TRC Rebuttal Document. TRC claims that the FFRA model significantly mis-predicts fluvial fish abundance at several sites that are selectively presented in the TRC report. As a follow-on to the USGS report, the FFRA model was extrapolated to the 1,429 subbasins in Massachusetts so that the average FFRA and 5 “biological categories” were defined based on August withdrawal intensity and % impervious cover in the subbasin (Biological Cateogires listed on p. 3 of TRC Report and “draft” statewide map of subbasin classification in Figure 5 of TRC Report). TRC compared archived fish data counts for stream localities within several (non-randomly) selected subbasins in different biological classes (TRC Report, Table 1 on p. 12) and compared these to the expected FFRA values provided in the USGS/MDFW Database. The TRC fish data were taken from several years spanning a range of flow conditions.

TRC evaluates these sites and concludes that the USGS/MDFW FFRA model fails to adequately predict the TRC-presented fish count data.

There is a serious **flaw** in this analysis having to do with using *incomparable units of the statistical response variable*.

The analysis presented in the TRC Rebuttal Document analyzes the fish <u>count</u> <u>data</u> (total number of captured fish) from the archived samples. The USGS report and FFRA model relies on a <u>standardized</u> <u>count</u>, i.e. the number of fish captured divided by the effort (CPUE), or time required to catch the fish. (CPUE is a standard sampling practice in fisheries biology.) Because the TRC data are not standardized by sampling time, the comparison presented in the TRC Document <u>cannot be meaningfully interpreted</u>. For example, if 100 fish were caught at a locality in 2 hours time, the USGS/MDFW value for this observation would be 50 fish (per hour), whereas the TRC value would be 100 fish. Clearly, lack of standardization can lead to apples-to-oranges type comparisons.

(5) The interannual variation in fish numbers at a handful of sites highlighted in the TRC Report undermines the utility of the USGS model to serve in a regulatory framework, because fish abundance did not vary consistently with interannual flow conditions

TRC also claims that because the case study fish data are sampled over several years of variable low flow conditions, the USGS/MDFW approach that assigns biological condition to overall flow alteration is unsupported.

However, again, the units of comparison preclude such inference because annual fish count data are not presented as CPUE. For example, if 200 fish were caught at site X in one year and 300 fish caught at the same site the next year but in three times the sampling time, the TRC analysis would report these as the same values (200 fish); however, the FFRA would report these as 300 in year 1 and 100 in year 2. These are large discrepancies based on lack of comparable units.

TRC presents fish count data from different years and in some cases different localities within each subbasin. This raises the question of how to select "representative" samples (streams) to "test" the application of the USGS/MDFW FFRA model (after standardizing the fish count data by effort), since no criteria appear to be presented in the USGS report or other supporting documentation I have seen. Presumably, this is to be worked out in further refinement of the Framework.

The inter-annual fluctuations in fish counts reported by TRC are instructive to the process of setting flow alteration standards. Stream fish are certainly adapted to survive periodic low flow conditions (or to recover from hydrologically connected flow refugia (such as deep pools or groundwater-fed tributary streams), and these low flows in the TRC report may well depress local abundances relative to more favorable summer flows. However, what the FFRA model "predicts" is that a prolonged (mutli-year, continuous) depression of late summer low flows is associated with reduced fluvial fish abundance (Figure 20A and Table 8), all other factors held constant. The TRC report does not evaluate variation in fish count data with respect to *flow alteration*, i.e., across subbasins having different degrees of August withdrawals. Rather, the TRC Report uses inter-annual variation in fish count data in response to natural flow variation to purportedly demonstrate that the USGS/MDFW model is "inconsistent" with observation. However, again, the RFFA model predicts *average* change in fish CPUE (fish/hour) given a *change* in August withdrawal. The definition of "average" is open for discussion at the subbasin level, since it includes many sites with potentially heterogeneous natural conditions across many years. In other words, as a statistical property, the "average" is comprised of many observations that can vary from place to place and year to year.

The interannual vairations within a particular site reported by TRC represent, in ecological parlance, a "pulse" experiment, where stressful conditions are occasionally imposed but the population has opportunities to recover in more favorable years. Recovery is enhanced where nearby refugia are accessible.

By contrast, persistent flow depletion from consistent late summer groundwater pumping each year would be the equivalent of a "press" experiment in ecology, i.e., where flow levels are continuously depressed over years and decades. One "test" of a press experiment would be to compare fluvial fish response across a gradient of flow alteration intensity. This is precisely what the USGS/MDFW FFRA models (both Quantile regression and GLM) do. The TRC Report does not speak to this directly because inter-basin comparisons along a flow alteration intensity gradient are not subject to analysis in the TRC Rebuttal document.

Another appropriate "test" of the FFRA's prediction would be to impose a multi-year experimental press experiment on an unregulated site to see if fluvial fish abundance declined over time in response to the AUGgwWp. This would not seem feasible.

Concluding Remarks:

Based on the information available to me, I find the TRC Report to have serious flaws as a science-based challenge to the USGS/MFDW FFRA model. All of the major assertions in the TRC Rebuttal Document are poorly supported and themselves rebutted by a careful reading of the USGS Report and a clarification of the non-commensurate units and scales of comparison TRC employs.

More generally, and to the issue of Original Question 3, after spending a fair amount of time contemplating the issues and methods at play, I am of the scientific opinion that the TRC Rebuttal Document does not afford a reasonable challenge to the use of the USGS/MDFW approach to guide policy development for water withdrawals in the State of Massachusetts.

The FFRA model and its associated Biological Condition Categories appear to be a sound science-based screening tool. That is to say, this approach is generally valid at the statewide (broad regional) level, based on the QR and GLM FFRA models, but the local application probably requires additional detail.

Importantly, even though the TRC Rebuttal Document documents interannual variation in fish counts (but <u>not</u> in CPUE, the appropriate but unmeasured metric) for a handful of non-randomly selected sites, this finding does not invalidate the inference from the FFRA model that flow alteration depresses fluvial fish abundance. That evidence is clearly provided in the quantile regression and GLM models.

Nonetheless, the TRC Report does expose some important issues in terms of how the USGS/MDFW FFRA Model might be used as a regulatory framework at specific localities.

The issue raised, specifically, is how to translate the FFRA model (created from at-a-site fish data and environmental variables) to a measure of Biological Condition defined at the whole subbasin scale (where there are many sites and potentially a range of at-a-site environmental conditions) (Figure 5 in the TRC Report).
The fact that the USGS/MDFW GLM model "explains" a total of 18% of total variation in FFRA across the state, and the Quantile Regression shows that there are many localities with low FFRA despite low levels of flow alteration. This reflects the importance of accounting for unmeasured environmental factors in more precisely predicting fluvial fish abundance at individual localities, not to mention whole subbasins.

This problem is not unique to the Massachusetts situation; rather, it is a challenge in any "scaling up" of site-specific data to the watershed scale. Specifically, the FFRA model is based on "local information" (i.e., at a locality) and within a subbasin, there are many localities which are likely to be variable in local conditions. For example, according to the FFRA model (Equation 6), wetland buffers and channel slope also influence RFFA, but these finer scale data are not captured in the Biological Condition Category definitions. Presence of geomorphic complexity (e.g., deep water or backwater refugia) or proximity to unregulated tributary refugia could also generate among-site variation in fluvial fish abundance within a subbasin. In other words, the subbasin Biological Category definitions represent an expected "average" condition, but within-subbasin heterogeneity in important environmental factors (wetlands, channel slope, refugia, connectivity, even distance from groundwater extraction sites) would be expected to generate some spatial variation in FFRA within a subbasin. Thus, in my view, the Biological Condition categories are useful as a first cut at capturing "mean condition" of a subbasin relative to other subbasins, but a finer scale (at a locality) information will reveal the extent to which a particular site may deviate from the average expectation.